An optimal superfarthingale and its convergence over a computable topological space

Kenshi Miyabe

Research Institute for Mathematical Sciences, Kyoto University

Abstract. We generalize the convergenece of an optimal semimeasure to a real probability in algorithmic probability by using game-theoretic probability theory and the theory of computable topology. Two lemmas in the proof give as corollary the existence of an optimal test and an optimal integral test, which are important from the point of view of algorithmic randomness. We only consider an SCT_3 space, where we can approximate the measure of an open set. Our proof of almost-sure convergence to the real probability by a superfarthingale indicates why the convergence in Martin-Löf sense does not hold.

1 Introduction

Algorithmic probability [11, 9] gives us a formal theory of inductive inference. The underlying space is the space of sequences over a finite alphabet. The main tool is a semimeasure. In this setting, Solomonoff's [18, 19] central result is explained in the following. Given a finite string, each semimeasure induces a subjective probability of the next alphabet. There exists an optimal semimeasure. Then the probability induced by the optimal semimeasure converges to the real computable probability almost surely.

In this paper, we make a first step to generalize algorithmic probability to a general space. The underlying space will be a space of sequences over a topological space. Our main results (Theorems 5 and 6) can informally described as follows. We want the probability that the next point falls in a subset of the space given a finite sequence of points. There exists an optimal function which induces the probability. Then the probability induced by the function converges to the real probability almost surely.

A semimeasure is essentially an equivalent notion for a supermartingale in algorithmic randomness [3, 14]. Algorithmic randomness over a general space is studied by some researchers [5, 8, 12]. The generalization uses computable topology studied by [21, 22]. To obtain the desired result, we require condition SCT_3, which is a computable separation axiom studied in [23]. To generalize a supermartingale in algorithmic randomness, Miyabe [12] uses a martingale in probability theory. In this paper, however, we use a superfarthingale [2, 20], which is a prequential version of a supermartingale in game-theoretic probability theory [17]. The effectivization of game-theoretic probability theory is a generalization of algorithmic probability. In this setting we prove the existence of an optimal superfarthingale. We also prove convergence to a real probability.

With lemmas developed in the proof of the existence of an optimal superfar-thingale, we prove Theorem 3 and Theorem 4 as corollaries. These are important from the point of view of algorithmic randomness.

Theorem 3 says that there exists an optimal test over an SCT_3 space. Hoyrup and Rojas [8, 7] proved the existence of an optimal test over a computable metric space and it is known that a computable metric space can be constructed from an SCT_3 space [6]. We give another more direct proof here.

Theorem 4 says that there exists an optimal integral test over an SCT_3 space with a computable measure. The existence of an optimal integral test over Cantor space is well-known [11]. Hoyrup and Rojas [8] proved the existence of an optimal integral test over a computable metric space with a computable probability measure. We will prove the existence of an optimal integral test over a SCT_3 space with a computable measure. Note in particular that our proof does not use the distance function.

In Section 3 we introduce a notion of approximation. Then we prove the existence of an optimal test. In Sections 4 and 5 we prove the existence of an optimal integral test and an optimal superfarthingale respectively and discuss when and to which measure the prediction converges.

2   Preliminaries

We recall relevant results from various fields, as some of the terminology and notations are not standard. Results in algorithmic probability and algorithmic randomness are for comparison to our results.

We assume that the readers are familiar with computability theory [15, 16] on natural numbers. A subset of $\mathbb{N}$ is *c.e.* if there exists a computable function such that the set is the domain of the function. A function $f : X \to \mathbb{R}$ is *c.e.* if $\{q \in \mathbb{Q} : q < f(x)\}$ is *c.e.* uniformly in $x \in X$. If X is countable, the meaning of the definition is clear. If X is a topological space, we refer the reader to Section 2.4 for details. A function $f : X \to \mathbb{R}$ is *computable* if f and $-f$ are c.e.

2.1   Algorithmic probability

For details see [11, 9]. Here we only consider binary sequences. The set of finite binary strings is denoted by 2^*. A *semimeasure* is a function $\mu : 2^* \to \mathbb{R}$ such that $\mu(\lambda) \leq 1$ and $\mu(\sigma) \geq \mu(\sigma 0) + \mu(\sigma 1)$ for all $\sigma \in 2^*$ where λ is the empty string. If the equalities hold, then μ is called a *probability measure*. A c.e. semimeasure μ_0 is *optimal*[1] if, for all c.e. semimeasure μ, there exists a constant $c > 0$ such that for all $\sigma \in 2^*$, we have $\mu_0(\sigma) \geq c \cdot \mu(\sigma)$. There exists an optimal c.e. semimeasure. For $\sigma \in 2^*$, σ_n is the n-th bit of σ, $\sigma_{t:n} = \sigma_t \sigma_{t+1} \cdots \sigma_{n-1} \sigma_n$ and $\sigma_{<n} = \sigma_{1:n-1}$. We also let $\mu(\sigma_n | \sigma_{<n}) = \mu(\sigma_{1:n}) / \mu(\sigma_{<n})$.

---

[1] The terminology "universal" is more common.

Theorem 1. *Let μ_0 be an optimal c.e. semimeasure and μ be a computable probability measure. Then*

$$\frac{\mu_0(\sigma_n|\sigma_{<n})}{\mu(\sigma_n|\sigma_{<n})} \to 1 \ \text{as} \ n \to \infty \qquad (1)$$

with μ-probability 1.

Informally speaking $\mu_0(\sigma_n|\sigma_{<n})$ is the prediction induced from μ_0. If the sequence is sampled by the measure space with computable μ, then the prediction converges to the real probability almost surely.

2.2   Algorithmic randomness

For details see [14, 3]. *Cantor space* 2^ω (the set of infinite sequences) is equipped with the topology generated by cylinders $[\![\sigma]\!] = \{A \in 2^\omega : \sigma \prec A\}$ where $\sigma \prec A$ means that σ is a prefix of A. An open set U on 2^ω is *c.e.* if $U = \bigcup\{[\![\sigma]\!] : \sigma \in S\}$ where S is a c.e. set of binary strings. Let μ be a measure on 2^ω.

A μ-*Martin-Löf test* is a sequence $\{U_n\}$ of uniformly c.e. open sets with $\mu(U_n) \le 2^{-n}$. A sequence $\alpha \in 2^\omega$ is μ-*Martin-Löf random* if $\alpha \notin \bigcup_n U_n$ for all μ-Martin-Löf tests. A μ-Martin-Löf test $\{U_n\}$ is *optimal* if, for any other μ-Martin-Löf test $\{V_n\}$, there exists c such that $V_{n+c} \subseteq U_n$ for all n. Let M be an optimal semimeasure. Then a sequence α is μ-Martin-Löf random iff there is a constant c such that $M(\alpha_{1:n}) \le c \cdot \mu(\alpha_{1:n})$ for all n. The set of μ-Martin-Löf random sequences has μ-measure 1. By Theorem 1 the set of sequences satisfying (1) also has μ-measure 1. Remarkably, the latter set does not contain the former set in general. We will consider this problem in Section 5.2.

Theorem 2 (Hutter and Muchnik [10]). *There exists an optimal semimeasure M and a computable measure μ and μ-Martin-Löf random sequence α such that*

$$M(\alpha_n|\alpha_{<n}) \not\to \mu(\alpha_n|\alpha_{<n}) \ \text{for} \ n \to \infty.$$

2.3   Game-theoretic probability

Game-theoretic probability theory [17] is influenced by Dawid's prequential principle [1], which says that our evaluation of the quality of the forecasts $p_1, p_2 \ldots$ in light of the observed outcomes $x_1, x_2 \ldots$ should not depend on Forecaster's model even if it exists and is known. In view of this a (super)farthingale was introduced in [2].

Let X be a topological space and $\mathcal{M}$ be a space of probability measures on X. Let $\Pi = (\mathcal{M} \times X)^\infty$ and $\Pi^\diamond = (\mathcal{M} \times X)^*$. For $\pi = (p_1, x_1, p_2, x_2, \ldots) \in \Pi$, let $\pi^n = (p_1, x_1, \ldots, p_n, x_n) \in \Pi^\diamond$.

Definition 1 (superfarthingale). *A* farthingale *is a function* $V : \Pi^\diamond \to [-\infty, \infty]$ *satisfying*

$$V(p_1, y_1, \ldots, p_{n-1}, y_{n-1}) = \int_X V(p_1, y_1, \ldots, p_{n-1}, y_{n-1}, p_n, x)dp_n \qquad (2)$$

for all n, all $(p_1, y_1, p_2, y_2, \ldots) \in \Pi$. If we replace "=" by "≥" in the equation (2), we get the definition of a superfarthingale.

If Forecaster behaves well, then a (super)farthingale V is bounded almost surely and a property stemming from the boundedness holds almost surely. Such a theorem can be easily converted to a theorem in measure-theoretic probability.

In [20], the case $\Omega = \{0, 1\}$ is considered and the existence of an optimal superfarthingale is shown. Furthermore a game-random sequence is defined as the sequence π such that $\sup_n V(\pi^n) < \infty$ for all c.e. superfarthingales V. The game-randomness is a generalization of Martin-Löf randomness. To define optimality of a superfarthingale, we need the definition that a superfarthingale is c.e.

2.4   Computable topology

We mainly refer to [22, 23] for computable topology. Let Σ^* and Σ^ω be the respective sets of the finite and infinite sequences, respectively, of symbols from a finite alphabet. A function mapping finite or infinite sequences of symbols from Σ is computable if it can be computed by a Type-2 machine. A representation of M is a surjective function $\gamma :\subseteq Y \to M$ where $Y \in \{\Sigma^*, \Sigma^\omega\}$. An object x is γ-computable if it has a computable representation p such that $\gamma(p) = x$.

A computable topological space is a 4-tuple $\mathbf{X} = (X, \tau, \beta, \nu)$ such that (X, τ) is a topological T_0-space, $\nu :\subseteq \Sigma^* \to \beta$ is a notation of a base β of τ, $\mathrm{dom}(\nu)$ is computable and $\nu(u) \cap \nu(w) = \bigcup\{\nu(w) : (u, v, w) \in S\}$ for all $u, v \in \mathrm{dom}(\nu)$ for some c.e. set $S \subseteq (\mathrm{dom}(\nu))^3$. For the points, the open sets and the closed sets, we use representations δ, θ and ψ^- that are defined as follows. For $p \in \Sigma^\omega$ and $x \in X$, $\delta(p) = x$ iff p is a list of all $u \in \mathrm{dom}(\nu)$ such that $x \in \nu(u)$, $\theta(p)$ is the union of all $\nu(u)$ where u is listed by p, and $\psi^-(p) = X \backslash \theta(p)$.

Computable separations are studied in [23] in detail but we use only SCT$_3$.

Definition 2 (SCT$_3$). *There are a c.e. set $R \subseteq \mathrm{dom}(\nu) \times \mathrm{dom}(\nu)$ and a computable function $r :\subseteq \Sigma^* \times \Sigma^* \to \Sigma^\omega$ such that for all $u, w \in \mathrm{dom}(\nu)$, $\nu(w) = \bigcup\{\nu(u) : (u, w) \in R\}$ and $(u, w) \in R \Rightarrow \nu(u) \subseteq \psi^- \circ r(u, w) \subseteq \nu(w)$.*

For computability of measures we refer to [13]. Let $\mathcal{M}(X)$ be the space of finite non-negative Borel measures on X. The following sets form a countable subbase of the A-topology on $\mathcal{M}(X)$: $\{\mu : \mu(G) > q\}$, $\{\mu : \mu(X) < q\}$ where G is the finite union of base sets and $q \in \mathbb{Q}$. Let β_A denote the base generated from the above subbase and ν_A denote a natural computable notation of the base β_A. Then the 4-tuple $\mathbf{M} = (\mathcal{M}(X), \tau_A, \beta_A, \nu_A)$ is a computable topological space. The representation δ of points in $\mathcal{M}(X)$ is denoted by δ_A.

3   Optimal test

In this section we give a proof of Lemma 1, which states an important property of a SCT$_3$ space with a computable measure and plays a central role in the proof of Theorem 5. As a corollary we obtain the existence of an optimal test over a SCT$_3$ space.

3.1 Approximation

Let $\mathbf{X} = (X, \tau, \beta, \nu)$ be a computable topological space. Let μ be a measure on $\mathbf{X}$. Note that μ need not be computable.

We give an approximation of the measure of an open set. An approximation is usually a sequence of rationals or sometimes a sequence of computable reals. However, the following approximation may be neither and carries a weaker notion.

Definition 3 (Approximation). *A μ-approximation of an open set $U \subseteq X$ is a sequence (U_n, a_n) such that U_n is the finite union of base sets, $U_n \uparrow U$, a_n is a rational, $\mu(U_n) \leq a_n$ and $a_n - \mu(U_n) \leq 2^{-n}$ for each n.*

As is usual we say that a μ-approximation (U_n, a_n) is $(\bigcup \nu^{\mathrm{fs}}, \rho)$-computable if U_n is uniformly $\bigcup \nu^{\mathrm{fs}}$-computable and a_n is uniformly ρ-computable.

To state lemmas and propositions clearly and concisely, we introduce the notion of γ-relativeness for a representation γ.

Definition 4 (Relative computability). *Let $\mathbf{X}_i = (X_i, \tau_i, \beta_i, \nu_i)$ be computable topological spaces for $i = 1, 2$. A point $x_2 \in X_2$ is δ_2-computable δ_1-relative to $x_1 \in X_1$ if there exists a (δ_1, δ_2)-computable function $f :\subseteq X_1 \to X_2$ such that $f(x_1) = x_2$.*

Relative computability of open sets, closed sets and functions are similarly defined.

Lemma 1. *Let $\mathbf{X}$ be a SCT_3 space. For an open set U we can uniformly construct a μ-approximation that is $(\bigcup \nu^{\mathrm{fs}}, \rho)$-computable (θ, δ_A)-relative to (U, μ).*

Proof. It suffices to show the lemma for a base set $\nu(w)$. Let R and r be the c.e. set and the computable function for SCT_3 from Definition 2. Let $\{u_i\}$ be a computable enumeration of the c.e. set $\{u \in \mathrm{dom}(\nu) : (u, w) \in R\}$.

For each m, let $U_m = \bigcup_{i=1}^m \nu(u_i)$ and $V_m = \bigcup_{i=1}^m \psi^- \circ r(u_i, w)$. Note that U_m is open and V_m is closed. Then $U_m \subseteq V_m \subseteq \nu(w)$ and $U_m \uparrow \nu(w)$. It follows that $\mu(U_m) \to \mu(\nu(w))$ and $\mu(V_m) - \nu(U_m) \to 0$ as $m \to \infty$.

Note that the measure $\mu(U_m)$ is approximated from below and $\mu(V_m)$ from above δ_A-relative to μ. Then we can compute m_n from a representation of μ such that $\mu(V_{m_n}) - \nu(U_{m_n}) \leq 2^{-n}$.

Let $\mu(U_{m_n})[s]$ and $\mu(V_{m_n})[s]$ be an approximation of rationals. Let s_n be the first stage such that $\mu(V_{m_n})[s_n] - \nu(U_{m_n})[s_n] \leq 2^{-n}$. Let $a_n = \mu(V_{m_n})[s_n]$. Then a_n is uniformly ρ-computable δ_A-relative to μ. Note that $\mu(U_{m_n})[s_n] \leq \mu(U_{m_n}) \leq \mu(V_{m_n}) \leq \mu(V_{m_n})[s_n] = a_n$. It follows that $a_n - \mu(U_{m_n}) \leq 2^{-n}$. Hence (U_{m_n}, a_n) is a μ-approximation. $\square$

3.2 Existence of an optimal test

Definition 5. *A μ-test over $\mathbf{X}$ is a sequence $\{V_n\}$ of uniformly θ-computable sets with $\mu(V_n) \leq 2^{-n}$. A μ-test $\{W_n\}$ is optimal if, for each μ-test $\{V_n\}$, there exists c such that $V_{n+c} \subseteq W_n$ for each n.*

Theorem 3. *Let* **X** *be a* SCT$_3$ *space and* μ *be a computable measure on* **X**. *Then there exists an optimal* μ-*test over* **X**.

Proof. There exists a computable enumeration $V_{m,n}$ of θ-computable sets. For each $V_{m,n}$, there exists a μ-approximation (U_s, a_s) by Lemma 1, which is uniformly $(\bigcup \nu^{\mathrm{fs}}, \rho)$-computable. Let $\widetilde{V}_{m,n} = U_{\sup\{s:a_s<2^{-n}\}}$. Then

$$\mu(\widetilde{V}_{m,n}) = \mu(U_{\sup\{s:a_s<2^{-n}\}}) = \sup\{a_s : a_s < 2^{-n}\} \leq 2^{-n}.$$

Hence $\{\widetilde{V}_{m,n}\}_n$ is a μ-test for each m.

Suppose $\mu(V_{m,n}) < 2^{-n}$. Then there exists s_0 such that $\mu(V_{m,n}) + 2^{-s} < 2^{-n}$ for all $s \geq s_0$. For such s we have $a_s \leq \mu(U_s) + 2^{-s} \leq \mu(V_{m,n}) + 2^{-s} < 2^{-n}$. Hence $\widetilde{V}_{m,n} = V_{m,n}$.

Let $W_n = \bigcup_m \widetilde{V}_{m,n+m+1}$. We claim that $\{W_n\}$ is an optimal μ-test. Note that $\mu(W_n) \leq \sum_m 2^{-n-m-1} \leq 2^{-n}$. Hence $\{W_n\}$ is a μ-test.

Let A_n be a μ-test. Then $\{B_n\}_n = \{A_{n+1}\}_n$ is also a μ-test with $\mu(B_n) \leq 2^{-n-1}$. Hence there exists m such that $B_n = V_{m,n}$. It follows that $W_n \supseteq V_{m,n+m+1} = B_{n+m+1} \supseteq A_{n+m+2}$ for each n. Then $\{W_n\}$ is optimal. $\square$

4   Optimal integral test

The goal in this section is to establish the existence of an optimal integral test over a SCT$_3$ space. The key of the proof is Lemma 2. As in Section 3, this lemma is also the key to proving Theorem 5 in the next section.

4.1   Computable bound

We gave an approximation of an open set in the previous subsection. By using this we give an approximation of a $(\delta, \overline{\rho}_<)$-continuous function.

Lemma 2. *Let* $f : X \to \overline{\mathbb{R}}^+$ *be a total function and* c *be a* ρ-*computable real. Then there exists a* $([\delta \to \overline{\rho}_<], \delta_A, [\delta \to \overline{\rho}_<])$-*computable function* b_c *such that (i)* $\int b_c(f, \mu)d\mu \leq c$, *(ii)* $\int f d\mu < c \Rightarrow b_c(f, \mu) = f$.

Proof. Let $\mathbb{Q}^+$ be the set of all non-negative rationals. Let $Q_n = \{q_i : i \leq n\}$ where $\{q_i\}$ is a computable enumeration of $\mathbb{Q}^+$. Then $Q_n \uparrow \mathbb{Q}^+$. Let $s(n,i)$ be the index of $\min\{r \in Q_n : r > q_i\}$. If q_i is the maximum element in Q_n and the set is empty, then $s(n,i) = \infty$. Then

$$\mu(f) = \int_X f d\mu = \lim_{n \to \infty} \sum_{i \leq n} q_i \cdot \mu(\{x : f(x) > q_i\} \backslash \{x : f(x) > q_{s(n,i)}\})$$

where $q_\infty = \infty$.

Let $U^i = \{x : f(x) > q_i\}$. Note that U^i is a θ-computable open set. Then it has a μ-approximation (U^i_n, a^i_n) satisfying the condition in Lemma 1. Note that there exists a computable function $b : \mathbb{N} \to \mathbb{N}$ such that $\sum_{i=0}^\infty q_i \cdot 2^{-b(i)} \leq 1$. Let

$$V^i_n = U^i_{n+b(i)} \backslash U^{s(n,i)}_{n+b(s(n,i))}.$$

Then $\mu(V_n^i) \to \mu(U^i \setminus U^{s(n,i)})$. Note that $\mu(f) \leq \sum_{i=0}^n q_i \cdot \mu(V_n^i)$ for each n and $\mu(f) = \lim_{n \to \infty} \sum_{i=0}^n q_i \cdot \mu(V_n^i)$.

Let $c(n,i) = a_{n+b(i)}^i - a_{n+b(s(n,i))}^{s(n,i)}$. Then $a_{n+b(i)}^i - 2^{-(n+b(i))} \leq \mu(U_{n+b(i)}^i) \leq a_{n+b(i)}^i$ for each n and i. It follows that

$$c(n,i) - 2^{-(n+b(i))} \leq \mu(V_n^i) \leq c(n,i) + 2^{-(n+b(s(n,i)))}.$$

We construct $b = b(f, \mu)$ as the limit of functions b_n. We define b_n inductively on n. Let $b_0(x) = 0$ for all x. If

$$\sum_{i \leq n} q_i \cdot (c(n,i) + 2^{-(n+b(s(n,i)))}) < c,$$

then let $b_n(x) = \max\{q_i : x \in U_{n+b(i)}^i\}$, otherwise $b_n = b_{n-1}$.

Since U_n^i is increasing on n, $b_n(x) \geq b_{n-1}(x)$ for all n and x. Note that

$$\int b_n d\mu = \sum_{i \leq n} q_i \cdot \mu(V_n^i) \leq \sum_{i \leq n} q_i \cdot (c(n,i) + 2^{-(n+b(s(n,i)))}) < c.$$

It follows that $\int b d\mu = \int \lim_n b_n d\mu = \lim_n \int b_n d\mu \leq c$.

Suppose $\int f d\mu < c$. Then $\int f d\mu + 2^{-n+1} < c$ for a large n. Then for such n,

$$c > \sum_{i \leq n} q_i \cdot \mu(V_n^i) + 2^{-n+1} \geq \sum_{i \leq n} q_i \cdot c(n,i) - 2^{-n} + 2^{-n+1}.$$

Combining the previous two results, we obtain

$$\sum_{i \leq n} q_i \cdot (c(n,i) + 2^{-(n+b(s(n,i)))}) \leq \sum_{i \leq n} q_i \cdot c(n,i) + 2^{-n} < c.$$

It follows that $b(x) = \lim_n b_n(x) = \sup\{q : f(x) > q\} = f(x)$ for all x.   □

4.2   Computable enumeration

Now it suffices to show that we have a computable enumeration of $(\delta_X, \overline{\rho}_<)$-computable functions.

Lemma 3. *There is a computable enumeration of $(\delta_X, \overline{\rho}_<)$-computable functions $f_n : X \to \overline{\mathbb{R}}^+$.*

To prove this, we use the following representation.

Definition 6 (δ_4-representation; see [23]). *Define a multi-representation $\overrightarrow{\delta_4}$ of the set $\mathrm{CP}(X_1, X_2)$ of all partial continuous functions $f :\subseteq X_1 \to X_2$ as follows:*

$$f \in \overrightarrow{\delta_4}(p) \iff \begin{cases} (w \ll p \Rightarrow (\exists u \in \mathrm{dom}(\nu_1), \ v \in \mathrm{dom}(\nu_2))w = \langle u, v \rangle) \\ and \ f^{-1}[\nu_2(v)] = \bigcup_{\langle u,v \rangle \ll p} \nu_1(u) \cap \mathrm{dom}(f). \end{cases}$$

Proof (Proof of Lemma 3). Let p be a $\overrightarrow{\delta_4}$-representation of a partial continuous function $f :\subseteq X \to \overline{\mathbb{R}}^+$. Then there exists only one total continuous function $f' : X \to \overline{\mathbb{R}}^+$ in $\overrightarrow{\delta_4}(p)$. Hence $\overrightarrow{\delta_4}$ is also a representation of the set $C(X, \overline{\mathbb{R}}^+)$ of all total continuous functions. The function represented by p is determined by the set $\{w \in \Sigma^* : w \ll p\}$. Then there is a computable enumeration of all computable $\overrightarrow{\delta_4}$-representations. $\square$

4.3   The existence of an optimal integral test

Definition 7. *A μ-integral test on* **X** *is a $(\delta, \overline{\rho}_<)$-computable function $f : X \to \overline{\mathbb{R}}^+$ with $\mu(f) \leq 1$. A μ-integral test f_0 is* optimal *if, for each μ-integral test f, there exists c such that $f(x) \leq c \cdot f_0(x)$ for all $x \in X$.*

Theorem 4. *Let* **X** *be a SCT_3 space. For an arbitrary measure μ on* **X***, there exists an optimal μ-integral test. Furthermore we can construct an optimal μ-integral test uniformly from δ_A-representation of μ.*

Proof (Proof of Theorem 4). By Lemma 3 there exists a computable enumeration $f_n : X \to \overline{\mathbb{R}}^+$ of $(\delta_X, \overline{\rho}_<)$-computable functions. Let $f = \sum_{n=0}^{\infty} 2^{-n-1} b_1(f_n, \mu)$ where b_1 is in Lemma 2. Then $\int b_1(f_n, \mu) d\mu \leq 1$ and $\int f d\mu \leq 1$. Hence f is an integral test.

If $\int f_n d\mu < 1$, then $b_1(f_n, \mu) = f_n$ and $f_n(x) \leq 2^n \cdot f(x)$ for all $x \in X$. Suppose $\int f_n d\mu = 1$. Then there exists m such that $f_n(x) = 2 f_m(x)$ for all x. It follows that $\int f_m d\mu = 1/2$ and $f_n(x) = 2 f_m(x) \leq 2^{m+1} \cdot f(x)$ for all $x \in X$. Hence f is optimal. $\square$

5   Optimal superfarthingale

5.1   Effectivization of game-theoretic probability

In the following we often use a $([\delta_A, [\delta \to \overline{\rho}_<], \delta]^*, \delta_A, \delta, \overline{\rho}_<)$-computable superfarthingale. For simplicity we call it a c.e. superfarthingale. Let $\mathcal{V}$ be the class of all non-negative c.e. superfarthingales V with $V(\Lambda) = 1$.

Definition 8. *A non-negative c.e. superfarthingale V is* optimal *if, for any superfarthingale $V' \in \mathcal{V}$, there exists a constant c such that, for any $\sigma \in \Pi^\diamond$, $cV(\sigma) \geq V'(\sigma)$.*

Theorem 5. *If* **X** *is a SCT_3 space, then there is an optimal superfarthingale.*

Definition 9. *A superfarthingale V is* strict *if the equation (2) in Definition 1 holds as a strict inequality ("$>$") for all n.*

Lemma 4. *For a non-negative superfarthingale V, there exists a strict superfarthingale V' such that $2V' \geq V$.*

Proof. Let $V'(\pi^n) = \frac{2^n}{2^{n+1}-1} \cdot V(\pi^n)$. Then $2V' > V$ and $\int_X V'(\pi^{n-1}, p_n, x) dp_n \leq \frac{2^n}{2^{n+1}-1} \cdot \frac{2^n-1}{2^n-1} \times \frac{2^{n-1}}{2^n-1} \cdot V(\pi^{n-1}) = \frac{2^{2n}-2^n}{2^{2n}-2^{n-1}} \times V'(\pi^{n-1}) < V'(\pi^{n-1})$. $\qquad\square$

Lemma 5. *There exists a computable enumeration $\{V_m\} \subset \mathcal{V}$ such that each non-negative c.e. strict superfarthingale with $V(\Lambda) = 1$ appears in the enumeration.*

Proof. Let $V'(\pi^n) = \sup_{q < V'(\pi^{n-1})} b_q(V(\pi^n), p_n)$ for all $n \geq 1$ and $V'(\Lambda) = 1$. Then $\int_X V'(\pi^{n-1}, p_n, x) dp_n \leq V'(\pi^{n-1})$. Hence $V' \in \mathcal{V}$.

Suppose that V is strict. We prove $V'(\pi^n) = V(\pi^n)$ inductively on n. Suppose that $V'(\pi^{n-1}) = V(\pi^{n-1})$. Then $\int_X V(\pi^{n-1}, p_n, x) dp_n < q \Rightarrow b_q(V(\pi_n), p_n) = V(\pi_n)$. Hence $\int_X V(\pi_{n-1}, p_n, x) dp_n < V'(\pi_{n-1}) \Rightarrow V'(\pi_n) = V(\pi_n)$. By the strictness of V and the assumption, $V'(\pi^n) = V(\pi^n)$. $\qquad\square$

Proof (Proof of Theorem 5). Let V_n be a computable enumeratation of strict superfarthingales. Let $V = 2^{-n-1} V_n$. We prove that V is optimal. Let V' be a superfarthingale. Then by Lemma 4 and Lemma 5 there exists n and c such that $cV_n \geq V'$. Then $c \cdot 2^n V \geq c \cdot 2^n 2^{-n} V_n \geq V'$. $\qquad\square$

5.2   Convergence to a measure

We will prove a generalized version of Theorem 1. Let ν_n be the measures induced by the optimal superfarthingale V, that is, $\nu_n(A) = \int_A \overline{V}(\pi^{n-1}, p_n, x) dp_n$ where $\overline{V}(\pi^n) = \frac{V(\pi^n)}{V(\pi^{n-1})}$. We discuss what is the limit of the measures $\nu_n(A)$.

Fix $\pi = (p_1, y_1, p_2, y_2, \ldots) \in \Pi$. Let W be a non-negative c.e. farthingale and μ_n be the measures induced by W. The Hellinger distance is defined by

$$h_n = \int_X \left(\sqrt{\overline{V}(\pi^{n-1}, p_n, x)} - \sqrt{\overline{W}(\pi^{n-1}, p_n, x)} \right)^2 dp_n$$

Let $N_n = \int_X \sqrt{\overline{V}(\pi^n)\overline{W}(\pi^n)} dp_n$. Then $N_n \leq 1 - \frac{h_n}{2} \leq \exp(-\frac{h_n}{2})$. The first inequality follows from $h_n = \int_X \overline{V}(\pi^n) dp_n + \int_X \overline{W}(\pi^n) dp_n - 2N_n \leq 2 - 2N_n$. The second inequality follows from $1 - x \leq e^{-x}$. Then $N_n \exp(\frac{h_n}{2}) \leq 1$.

There exists a superfarthingale Y such that

$$Y(\mu_1, y_1, \mu_2, y_2, \ldots) = \sqrt{\frac{V(\pi^n)}{W(\pi^n)}} \cdot \exp(\frac{1}{2} \sum_{i=i}^{n} h_i),$$

because $\frac{\int_X Y(\pi^n) d\mu_n}{Y(\pi^{n-1})} = \int \sqrt{\overline{V}(\pi^n)\overline{W}(\pi^n)} dp_n \exp(\frac{h_n}{2}) = N_n \exp(\frac{h_n}{2}) \leq 1$.

Theorem 6. *If $\sup_n Y < \infty$ then $h_n \to 0$ as $n \to \infty$.*

Proof. By optimality of Y we have $\sum_{i=1}^{\infty} h_i < \infty$. Recall h_i is non-negative. $\quad\square$

Hence we conclude that $h_n \to 0$ as $n \to \infty$ almost surely. By Theorem 2, Y is not c.e. in general. This is because h_i is not. To obtain the convergence on a random sequence, we need a stronger randomness notion such as 2-randomness or difference randomness [4].

References

1. Dawid, A.P.: Statistical theory: the prequential approach (with discussion). Journal of the Royal Statistical Society A 147, 278–292 (1984)
2. Dawid, A.P., Vovk, V.: Prequential probability: principles and properties. Bernoulli 5, 125–162 (1999)
3. Downey, R., Hirschfeldt, D.R.: Algorithmic Randomness and Complexity. Springer, Berlin (2010)
4. Franklin, J.N.Y., Ng, K.M.: Difference randomness. In: Proc. Amer. Math. Soc. vol. 139, pp. 345–360 (2011)
5. Gács, P.: Uniform test of algorithmic randomness over a general space. Theoretical Computer Science 341, 91–137 (2005)
6. Grubba, T., Schröder, M., Weihrauch, K.: Computable metrization. Mathematical Logic Quaterly 53(4-5), 381–395 (2007)
7. Hoyrup, M., Rojas, C.: An Application of Martin-Löf Randomness to Effective Probability Theory. In: CiE. pp. 260–269 (2009)
8. Hoyrup, M., Rojas, C.: Computability of probability measures and Martin-Löf randomness over metric spaces. Information and Computation 207(7), 830–847 (2009)
9. Hutter, M.: Universal artificial intelligence: Sequential decisions based on algorithmic probability. Springer (2005)
10. Hutter, M., Muchnik, A.: On semimeasures predicting Martin-Löf random sequences. Theoretical Computer Science 382, 247–261 (2007)
11. Li, M., Vitányi, P.: An introduction to Kolmogorov complexity and its applications. Graduate Texts in Computer Science, Springer-Verlag, New York, third edition. (2009)
12. Miyabe, K.: Algorithmic randomness over general spaces, submitted
13. Miyabe, K.: A computable topological space of measures, submitted
14. Nies, A.: Computability and Randomness. Oxford University Press, USA (2009)
15. Odifreddi, P.: Classical Recursion Theory, vol. 1. North-Holland (1990)
16. Odifreddi, P.: Classical Recursion Theory, vol. 2. North-Holland (1999)
17. Shafer, G., Vovk, V.: Probability and Finance: It's Only a Game! Wiley (2001)
18. Solomonoff, R.J.: A formal theory of inductive inference I, II. Information and Control 7, 1–22,224–254 (1964)
19. Solomonoff, R.J.: Complexity-based induction systems: Comparisons and convergence theorems. IEEE Transaction on Information Theory IT-24, 422–432 (1978)
20. Vovk, V., Shen, A.: Prequential randomness and probability. Theoretical Computer Science 411, 2632–2646 (2010)
21. Weihrauch, K.: Computable Analysis: an introduction. Springer, Berlin (2000)
22. Weihrauch, K., Grubba, T.: Elementary Computable Topology. Journal of Universal Computer Science 15(6), 1381–1422 (2009)
23. Weihrauch, K.: Computable Separation in Topology, from T_0 to T_3. In: CCA (2009)